DOCUMENT RESUME

ED 150 175 TM 006 890

AUTHOR Northrop, Lois C.
TITLE The Definition and Measurement of Judgment.
INSTITUTION Civil Service Commission, Washington, D.C. Personnel Measurement Research and Development Center.
REPORT NO PRDC-TM-76-17
PUB DATE Nov 76
NOTE 22p.; For related document, see TM 006 891

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Bibliographies; Cognitive Ability; Cognitive Measurement; *Decision Making Skills; *Factor Analysis; Literature Reviews; Measurement Techniques; Problem Solving; *Psychological Tests; Testing Problems; *Test Items
IDENTIFIERS Air Force; *Judgment

ABSTRACT

The complex nature of the psychological construct called judgment, the history of its identification as a mental ability, its measurement, and its relationship to other mental abilities are discussed. Factor analytic research which isolated the judgment factor is reviewed. The pervasiveness of the judgment factor throughout a number of diverse ability clusters (planning, problem solving, semantic evaluation, reasoning) make factor analysis particularly difficult. Sample test items measuring judgment ability and bibliographical references are appended. (Author/BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Personnel
Research and
Development
Center

Technical Memorandum 76-17

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED150175

The Definition and Measurement of Judgment

TM006 890

United States
Civil Service Commission
Bureau of Policies and Standards

Technical Memorandum 76-17

THE DEFINITION AND MEASUREMENT OF JUDGMENT

Lois C. Northrop

Test Services Section
Personnel Research and Development Center
U.S. Civil Service Commission
Washington, D.C. 20415
November 1976

THE DEFINITION AND MEASUREMENT OF JUDGMENT

ABSTRACT

The complex nature of the psychological construct called judgment, the history of its identification as a mental ability, its measurement, and its relationship to other mental abilities are presented and discussed in this paper. Factor analytic research which isolated the Judgment factor is reviewed and instruments used to measure it are examined.

CONTENTS

	Page
Introduction	1
The Nature of Judgment	1
Judgment as a Measurable Ability	1
The Limitations of Factor Analysis	2
The History of the Judgment Factor	3
Conclusions	6
Appendix	7
Reference Note	10
References	10
Annotated Bibliography	12

THE DEFINITION AND MEASUREMENT OF JUDGMENT

This paper reviews the history of the identification and clarification of the psychological construct called "judgment" and discusses its complex nature as well as the problems encountered in measuring it. This construct is one of the cognitive abilities commonly measured in written tests for personnel selection; for many years judgment has been measured in tests developed and used by the U.S. Civil Service Commission. The Commission's operational definition of this ability construct is as follows:

> "...the ability to solve a presented problem when all the necessary facts to solve the problem are not given. Solution to the problem involves making some reasonable assumptions or anticipating wh t the most likely of several possible occurrences might be. This ability becomes especially important when decisions must be reached based on incomplete evidence." (U.S. Civil Service Commission, 1974, p. 24.)

The Nature of Judgment

The concept of judgment used in such an operational definition is complex. The term judgment in the history of psychological testing has elicited a fair amount of controversy as to its exact nature and role in the structure of cognitive abilities. A number of discussions and other operational definitions of judgment have appeared in the literature. Hollingsworth (1925) defines judgment as a process involving relations wherein the individual (a) analyzes the idea (recognizing and comparing relationships within it), (b) synthesizes information and feelings with a resulting yes-no consciousness and (c) finally expresses a verdict. On the other hand, Johnson (1945, 1955) refers to judgment as the end product of several forms of intellectual activity (including analysis and synthesis). Material from perceptual, conceptual, recall and affective processes is evaluated in the act of judging and a decision is made which resolves an unsettled situation. This is consistent with Dewey's (1933) problem-solving model in which the final phase consists of evaluating possible solutions before employing one of them. However, evaluation may also be called for during the problem-solving process whenever information recognized, remembered or produced is tested for relevancy or value.

Whether judgment is defined as an end product only, or as an entire process, it cannot take place without the involvement of a number of successive mental processes. Relevant data must be abstracted from a situation and maintained in the mind; each piece of information must be weighted appropriately and finally integrated to relate to a chosen alternative. Judgment contains other aspects which enhance its complexity. Individual differences exist in how judgments are made. Whereas one person may accomplish the task via a process of successive hurdles (if the answer to question 1 is "yes", question 2 will follow; if the answer to question 1 is "no", a decision will have been made relative to an alternative), another may base a decision on general impression, and still a third may predicate a judgment on an addition of pros and cons. Judgment also calls for flexibility, because data can be examined from several points of view; different types of thinking (theoretical vs practical) can be involved; and the influence of training, experience, knowledge, the importance of the particular evaluation to the individual (ego involvement), and the speed with which judgments are made are all factors which must be considered.

Judgment as a Measurable Ability

Factor analytic studies to identify mental abilities have not consistently isolated a factor which could definitely be interpreted as Judgment. Thurstone's 1938 monograph on primary mental abilities did not report a Judgment factor but Corter (1952) has noted that Thurstone's factors were all highly object ve and somewhat barren as descriptions of human behavior. Cattell (1971) considers judgment to be one of the narrower, less substantiated mental abilities whereas Guilford (Guilford & Hoepfner, 1971) interprets judgment as a broad factor applying to a whole category of evaluative abilities - evaluation being the process of comparing items of information in terms of known specifications and reaching a decision which satisfies some criterion (either of similarity, identity or consistency). The term "evaluation", with its broader connotation, has been substituted for the term "judgment" in Guilford's research. Some authors (Johnson, 1955; Vernon, 1950)

believe that judgment belongs in the domain of reasoning but when the concept of reasoning is examined we discover that it is not uniquely defined.

General reasoning has been conceptualized as a broad property contained in many tests. It involves thinking activities directed at perceiving and solving unfamiliar problems - these naturally would include an analysis of ideas and facts, a weighing of inferences, implications, and possible consequences, and an ultimate organization of a sequence of related steps and integration of thought in order to formulate a conclusion. In addition to general reasoning, Guilford, Green and Christensen (1951) have substantiated two other "logical" reasoning factors - Induction (the ability to see underlying rules, trends and relationships or to reason from the specific to the general) and Deduction (the ability to make inferences or to reason from the general to the specific, especially characterized by problems of a syllogistic nature). French (1951) cites 37 other studies in which a Deduction (or Deductive Reasoning) factor has appeared and 9 other studies which have extracted an Induction factor. The identification of either of these factors is sometimes difficult because the kind of reasoning process used can vary with the type of problem, the respondent's approach to it, and his or her familiarity with other problems of a similar nature.

Acts of judgment or evaluation are necessarily a part of all of these intellectual functions (general reasoning, induction and deduction). The pervasive properties of both reasoning and judgment seem obvious; indeed Pawlik (1966) has classified judgment as a generalized solution mechanism, a technique which aids people in coping with their perceptual limits.

The question now arises as to how a multidimensional concept such as judgment can be measured and identified objectively. As has been mentioned, attempts to isolate a Judgment factor from among other mental abilities have not always been successful, but evidence does exist which indicates that although confounded with and related to planning abilities, creativity, verbal comprehension and reasoning, it is an important and assessable aspect of human intellect, necessary to practical success in numerous occupations (McKillip, Trattner, Corts, & Wing, Note 1).

The Limitation of Factor Analysis

Before proceeding with a review of specific factor-analytic studies which have demonstrated a Judgment factor, a brief discussion of the limits of the factor-analytic technique is presented to assist the reader in the interpretation and evaluation of these studies.

Abilities (as represented by extracted factors) are seldom, if ever, pure because personality traits, interests, motivation, relevant experience and situational conditions affecting performance all contaminate the variance contributed by ability. Furthermore factor analyses from several studies are not directly comparable because the number and kinds of tests used in each study are not alike and the populations used in testing vary as to background, age, sex and range of general aptitude. Only a few cognitive factors (Verbal, Numerical, Spatial) are very stable; i.e., they are consistently identified regardless of variations in population and test characteristics (Pawlik, 1966). Others are more dependent upon the particular selection of tests used and upon the nature of the population involved. It is essential to keep in mind that a factor analysis quantitatively clusters data which have aspects in common; the psychological interpretation of factors is left to the experimenter, who examines and appraises the content of the tests loading significantly together and arrives at a descriptive name for each cluster.

A further consideration when evaluating any factor analytic study is the total number of factors which have been extracted and rotated and the type of rotation (orthogonal or oblique) which has been employed. Decisions concerning these problems are guided either by information contained in the data itself (test scores and their intercorrelations), or by hypothesizing the number and character of expected factors in a certain area of performance and then constructing an adequate number of tests to represent each hypothesized factor. The latter method has been utilized by Guilford and his coworkers, many of whose studies will be cited in the pages to follow. Pawlik

(1966) has criticized Guilford's approach on the grounds that it classifies behavior by inspection rather than correlation. Wechsler (1958) points out that the profusion of factors identified in Guilford's work seems to contradict the purpose of factor analysis - i.e., to account for the variance of a large number of tests in terms of a minimal number of primary abilities. This criticism may or may not be valid depending upon how specific one wishes to make the categorization of ability. The subject is still controversial. The fact remains that when a large number of factors are obtained from an analysis, many of them will account for only 1 - 2% of the total variance and therefore each will account for a very specific aspect of behavior. Similarly, a factor loading of .2 (which is frequently considered significant when naming a factor) accounts for only 4% of a test's variance. Thurstone (1938) feels that only tests with factor loadings of .4 (16% of the total variance) should be included in the interpretation of a factor and that to name a factor with any real confidence, loadings of .5 or more are necessary. The number of tests contributing variance to a factor is also of importance; a factor named on the basis of but one or two tests is probably a very specific one, whose value will, again, depend upon the specificity of ability classification desired. There have been many regroupings and subsequent renaming of factors as a result of the different conditions and objectives of individual studies.

Lastly, there exists some evidence, based on cross-sectional rather than longitudinal studies, that one ability factor can be substituted for another in the mental organization of different age groups. Differentiation and specialization of ability is hypothesized up to the age of 25 or 30; thereafter, a generalization and integration of intellectual skills becomes effective and the older person is more likely to use several abilities in solving a single problem (Balinsky, 1941).

The History of the Judgment Factor

During World War II the Army Air Force carried on an extensive research program concerned with the development of instruments for the selection and classification of air crew personnel (Guilford & Lacey, 1947; French, 1951). Psychological constructs in the intellectual area were drawn up after a survey was made of instructors as to the reasons students failed in pilot training. Among the qualities considered essential by the instructors for successful pilots was "good judgment."

A number of test batteries were constructed - one contained Judgment and Reasoning tests and two others consisted of tests of Foresight and Planning. Tests of Judgment (which were included in all three batteries) were composed of items presenting a predicament and requiring the selection of a best solution. Practical (or commonsense) Judgment items presented problems of a "common everyday" type, some having to do with work planning, the solution to which rested not upon logical reasoning grounds but on the ability of the examinee to draw upon his "common sense," experience and general information background.[1] In addition to using given facts, additional reasonable assumptions had to be made. "Logical Reasoning Judgment" items were similarly constructed but were intended to measure the ability to draw a conclusion logically (needless to say these items loaded heavily on the factor Deduction as well as Judgment). Factor analyses of these Judgment tests showed that they definitely contained variance not accounted for by other well known factors and this was identified as a Judgment factor. Loadings ranged from .27 to .40 and additionally, an appreciable verbal component was present (loadings ranging from .26 to .46).[2] The Foresight and Planning Battery also contained a measure called Sequence of Maneuvers which asked the examinee to arrange a series of airplane maneuvers in a manner such that the least amount of climbing and diving would be required. In addition to the Judgment

[1]Much earlier than this the Army Alpha Examination contained a section purported to measure commonsense, wherein the examinee was to choose the best answer to a question - e.g., Why is wheat better for food than corn? Because:

a. it is more nutritious
b. it is more expensive
c. it can be ground finer

[2]Examples of items from a number of the Air Force tests are contained in the Appendix.

factor (a loading of .38), it loaded on the Numerical, Verbal and Planning factors.

The reliabilities of Judgment tests were uniformly low (.39 for Logical Reasoning Judgment and .23 for Commonsense Judgment). The communalities (variance in common with other tests) of the various forms of Judgment tests were generally not high (averaging .36), indicating either factorial complexity beyond those factors extracted in the several studies or considerable variance specific to each individual test. For judgment tests the former is undoubtedly the case since it has been established that the concept is a multidimensional one. As Adkins and Lyerly (1952) have pointed out, factorially pure tests of any kind are virtually impossible to achieve since the medium of the test (words, figures, numbers) intrudes. Although unnecessary test complexity should be avoided, items measuring judgment cannot be made so simple that the Judgment component is lost.

A re-analysis of some of the Air Force intercorrelation data by Adkins and Lyerly (1952) failed to identify Judgment as a factor. The Air Force test, Practical Judgment, achieved loadings of .25 on factors called Verbal, Visualization, and Non-Verbal-Non-Numerical Reasoning, each of which was associated with four to seven tests with factor loadings greater than .30. The sample sizes were not the same for all intercorrelations used, however, lending a note of skepticism to the results. Furthermore, a sufficient number and variety of tests related to judgment were not included (Practical Judgment was the sole example used in a total battery of 38 tests) so it is hardly surprising that no Judgment factor emerged.[3]

Similarly, items from the Air Force Practical Judgment test were used to construct an instrument called Practical Situations, which was included with 65 other variables in a battery of Reasoning Tests (Adkins & Lyerly, 1952). In the resultant factor analysis Practical Situations was found to have significant Verbal Relations, Spatial and Induction components, providing some documentation for the relationship of judgment and reasoning abilities. In a study of reasoning abilities by Corter (1952) a Judgment factor was also extracted. Tests loading .4 or greater on this factor included a modification of the Stanford-Binet Vocabulary test, a Reasoning test, Practical Judgment, Word Sorting, and the SRA Primary Mental Abilities Verbal test. Inductive as well as deductive processes seemed to be involved. A relationship between judgment and maturity, wisdom and experience was proposed as a possibility. Whereas Corter's subjects were 100 high school students, Guilford, Green, Christensen, Hertzka, and Kettner (1952, 1954) examined reasoning abilities using over 800 air cadets and student officers. Again a well-defined but more generalized Judgment factor emerged, identified by measures of Reading Comprehension, Inference, and Verbal Similarities as well as the Air Force Practical Judgment test. All of these tests seemed to satisfy the criteria for defining judgment - the evaluation of implications, the weighing of alternatives and the necessity for making assumptions beyond the given facts.

The nature of judgment was further explored by Hertzka, Guilford, Christensen, and Berger (1954) who, in a study of evaluative abilities, identified a Redefinition factor (the ability to shift the function of a common object and use it in a different way, sometimes referred to as a Gestalt transformation) which seemed to relate to judgment (assumptions beyond given facts required). Seeing Deficiencies was a measure which went significantly with this factor. Items for this test presented a practical judgment-type problem and a solution; the examinee must determine the chief fault with the given solution. In an analysis of reasoning, creative and evaluative abilities carried out at the University of Southern California Psychological Laboratory (Guilford, Kettner, & Christensen, 1956) the Gestalt transformation items and Seeing Deficiencies, as well as tests involving word associations and remote verbal similarities were grouped and identified as Judgment. In 1959 Kettner, Guilford and Christensen raised the question "Are Judgment and

[3]The usual practice of factor analytic investigations is to include several reference tests for a particular ability so that if its factor is present, it will emerge. At least three tests, and preferably more, are usually needed to identify a factor.

Redefinition really separate factors?" The issue was not resolved.

Judgment was isolated again by Guilford, Berger and Christensen (1955) and Berger, Guilford and Christensen (1957) in investigations of Planning Abilities. A possible relation between Judgment (work planning type problems) and Prediction (a part of planning) was hypothesized in the first study. Seeing Deficiencies was again identified with Judgment, its factor loading of .46 was somewhat higher than that of Practical Judgment (.30). The results of the latter study were similar, with the additional notation that a Judgment component was found in Lorge's General Planning Skills Test.

The foregoing review of factor analytic studies is plainly indicative of the complexity and pervasiveness of the Judgment concept. Further work was carried out by Guilford and his students but as the three-dimensional structure of intellect model developed, differentiating abilities more and more specifically, it is not surprising that Judgment, as broadly defined previously, disappeared as a factor. According to Guilford (1966), evaluation (a more inclusive, more definable term than judgment) is one of the five major operations (major intellectual processes) through which an individual deals with the "raw materials" of information. The others are cognition (discovery, comprehension and awareness), memory (retention, storage), divergent production (variety and quantity of output), and convergent production (reduction of information to achieve a best outcome). Evaluation itself involves reaching decisions or making judgments in order to satisy some criterion such as correctness, suitability or desirability. Each operation can be performed on one of four broad types of information (figural, symbolic, semantic and behavioral) and during processing, each of these types may take one of five forms - units (which have a "thing" character), classes (sets of items having common properties), relations (connections), systems (complexes of interacting parts), transformations (changes in information or its function), and implications (extensions in the form of expectancies, predictions or consequences).

In the initial study of evaluative abilities (Guilford, Hertzka & Christensen, 1953) a Judgment factor was not found. The Air Force Practical Judgment test achieved its highest and only significant loading (.42) on a General Reasoning factor (defined here as the ability to describe a problem and structure it in order to find a solution) which was also identified by tests of Arithmetic Reasoning, Reading Comprehension and Interpretations (essentially inference). Again we have substantiation for a relationship between judgment and general reasoning and the question again arises "Can judgment and general reasoning be separated?" Some years later Nihira, Guilford, Hoepfner, and Merrifield (1964) hypothesized a Judgment factor in a study of the semantic evaluative abilities (using over 400 high school students as subjects) but none was found. Judgment, as it has usually been conceptualized, seemed to be too broad to fit into Guilford's Structure of Intellect model. Two reference tests were employed, referred to as Commonsense Judgment I and II.[4] The former was in essence the Seeing Deficiencies test which had been used in several previously reported studies; the latter contained items of a similar structure and practical nature but the examinee was instructed to select the best methods (two out of five) of demonstrating the truth of a given statement. It was presumed that these tests would appear in the Semantic Transformations factor (probably because of Judgment's earlier association with Redefinition) but this was not so. Commonsense Judgment I was present in Evaluation of Semantic Classes (loading = .36), Relations (.30), Systems (.32), and Implications (.36). The measures Best Class Name (for a group of words), Language I.Q., Complete Thoughts, Logical Reasoning, and Verbal Analogies also appeared as contributors to these factors. In addition, Commonsense Judgment I correlated greater than .30 with Complete Thoughts, Verbal Analogies, and Logical Reasoning. Commonsense Judgment II, on the other hand, neither loaded on any of the Evaluative factors nor correlated significantly with any of the other tests in the battery.

Considering the nature of the structure of intellect model with its multitude of highly specific factors, the absence of a Judgment factor is not too disturbing. Judgment is a quality or ability necessary

[4]Examples of items appear in the Appendix.

to the solution of any complex problem. Its global character is evident; therefore it is unreasonable to attempt its assessment by any so-called "pure" measure. Any test of judgment will necessarily be contaminated with the many other abilities with which judgment is associated. In constructing such a test, one can only strive to minimize concurrent measurement of these other factors.

Conclusions

Perhaps the most salient impression one gets from the foregoing review of judgment as a measurable construct is of its pervasiveness throughout a number of diverse ability clusters - planning, problem solving, semantic evaluation, reasoning. Acts of judgment seem inevitable at some point in any endeavor requiring a decision, be it mechanically, creatively, artistically, theorectically, experimentally, or practically oriented. Obviously, factor pure measures of judgment are impossible, both because of its global nature and because the medium of a test (figural, symbolic, verbal) will necessarily intrude upon the measure of any ability.

APPENDIX

The following items are examples of those appearing in some of the Judgment tests referred to in this report.

1. Practical Judgment (nonmechanical), Guilford and Lacey, 1947, p. 128.

The principal reason why barracks in Army camps are built according to an identical plan is that this method:

A. Requires the least use of construction materials.
B. Makes possible the greatest speed of construction.
C. Allows construction of barracks that will last a long time.
D. Requires a low cost of upkeep of the barracks after completion.
E. Results in a military appearance which is similar in all Army camps.

Answer not given.

2. Practical Judgment (Work Planning type), Guilford and Lacey, 1947, p. 129.

At a mobile army encampment, it is necessary to use buckets to fight a fire 100 feet away from a stream and up a hill. Fifteen 5-gallon buckets and 60 men are available. The best procedure would be to:

A. Select 15 men and have each run between the stream and the fire carrying buckets and replace the entire crew at intervals.
B. Line up the 60 men from the building to the stream and pass the buckets from man to man from the stream to the building, the last man throwing the water on the fire and returning with the empty bucket.
C. Make one line of 25 men to pass the buckets from the stream to the building with one line of 10 men to throw the empty buckets back: 2 men to dip, 2 men to throw water on the fire, and the rest for relief.
D. Detail 2 men to dip and 2 men to throw water on the fire; assign 26 to carry buckets from the stream to the building and back; replace the entire crew with 30 fresh men in 15 minutes.
E. Detail 5 men to fill the buckets at the stream and leave them at the bank, where they can be picked up by the rest of the men who will carry them to building and return with empty buckets.

Answer not given.

3. Commonsense Judgment, Guilford and Lacey, 1947, p. 143.

A bomber squadron is over enemy territory on its way to bomb an oil refinery when one of the observes notices an advanced enemy airdrome. He notifies the squadron leader. It would be best for the squadron leader to:

A. Order the bombers in the squadron to continue as planned and report the location of the enemy airdrome on arrival at their base.
B. Order the squadron to circle the airdrome while he radios his position to his base.
C. Order half the bombers in the squadron to bomb the enemy airdrome and the other half to carry out the mission against the oil refinery.
D. Order all the planes in the squadron to bomb the enemy airdrome and return to their base.
E. Order the squadron to continue as planned while he returns to his base to report the location of the enemy airdrome.

Answer not given.

4. Logical Reasoning Judgment, Guilford and Lacey, 1947, p. 144.

An officer is in command of an advance unit in a night foray against distant enemy lines to test their strength in preparation for an assault against them. His orders are to send up a red flare if the enemy is not prepared for an attack, a blue flare if the enemy is prepared for an attack, a red flare and then a blue flare if the enemy is preparing an attack of its own, and a blue flare and then a red flare if the enemy is already beginning an attack. The officer finds the enemy beginning an attack, but instead of sending first a blue flare and then a red one, by mistake he sends up a red flare first. If he realizes his mistake immediately, it would be best for him to:

A. Wait several minutes and then send up a blue flare with a red flare following immediately after it.
B. Send up a blue flare right away and then a red flare immediately after it.
C. Send up a blue flare, wait a minute, and then send up a blue flare with a red flare immediately following it.
D. Send up a blue flare, wait a minute and then send up a red flare with a blue flare immediately following it.
E. Send a man back to the base to report the mistake.

Answer not given.

These four items are from Army Air Force Tests and are, of course, militarily oriented. The following items are from tests used by Guilford and his associates in various investigations.

5. Commonsense Judgment I (Seeing Deficiencies), Guilford and Hoepfner, 1971, p. 383.

Select the two best of five given reasons why a briefly described plan is faulty.

A city wants to improve its streets and its sewer system. It decides to work on street improvements first.

A. The streets will have to be torn up again in order to work on the sewer system.
B. Street improvements can be done gradually, so might as well be done second.
C. Since street improvement is less important than sanitation, the sewer system should be installed first.
D. Since the street improvement is costly, the city might run out of funds before it can work on the sewer.
E. The homeowners would disagree with the council that street improvement comes before sewer improvement.

Answers: A and C

6. Commonsense Judgment II Guilford and Hoepfner, 1971, p. 383.

Select the two best of five given methods to demonstrate the truth of a given statement.

The light in an electric refrigerator goes out the moment the door is closed.

A. Feel whether the bulb is hot after the refrigerator door has been closed for some time.
B. Substitute a wire-socket arrangement leading to a bulb outside the refrigerator.

C. Utilize a delayed shutter camera that would expose a film after the door is closed.
D. Disconnect all the appliances except the refrigerator, close the door, then see whether the electric meter is running.
E. Cut a hole in the refrigerator door.

Answers: B and C

REFERENCE NOTE

1. McKillip, R., Trattner, M., Corts, D., & Wing, H. Professional and Administrative Career Examination: Research and Development. Manuscript in preparation, Washington, D.C.: U.S. Civil Service Commission, 1976.

REFERENCES

Adkins, D.C., & Lyerly, S.B. Factor analysis of reasoning tests. Chapel Hill, N.C.: University of North Carolina Press, 1952.

Balinski, B. An analysis of the mental factors of various age groups from nine to sixty. Genetic Psychology Monographs, 1941, 23, 191-234.

Berger, R.M., Guilfrod, J.P., & Christensen, P.R. A factor-analytic study of planning abilities. Psychological Monographs, 1957, 71, No. 435.

Cattell, R.B. Abilities: their structure, growth and action. Boston: Houghton-Mifflin, 1971.

Corter, H.M. Factor analysis of some reasoning tests. Psychological Monographs, 1952, 66, No. 340.

Dewey, J. How we think. New York: D.C. Heath, 1933.

French, J.W. The description of aptitude and achievement tests in terms of rotated factors. Psychometric Monographs, 5, 1951.

Guilford, J.P. Structure of intellect factors and their tests. University of Southern California Psychological Laboratory Report 36, 1966.

Guilford, J.P., Berger, R.M., & Christensen, P.R. A factor-analytic study of planning. II. Administration of tests and analysis of results. University of Southern California Psychological Laboratory Report 12, 1955.

Guilford, J.P., Green, R.F., & Christensen, PR. A factor-analytic study of reasoning abilities. II. Administration of tests and analysis of results. University of California Psychological Laboratory Report 3, 1951.

Guilford, J.P., Green, R.F., Christensen, P.R., Hertzka, A.F., & Kettner, N.W. A factor-analytic study of Navy reasoning tests with the Air Force Aircrew Classification Battery. University of Southern California Psychological Laboratory Report 6, 1952.

Guilford, J.P., Green, R.F., Christensen, P.R., Hertzka, A.F., & Kettner, N.W. A factor-analytic study of Navy reasoning tests with the air Force Aircrew Classification Battery. Educational and Psychological Measurement, 1954, 14, 301-325.

Guilford, J.P., Hertzka, A.F., & Christensen, P.R. A factor-analytic study of evaluative abilities. II. Administration of tests and analysis of results. University of Southern California Psychological Laboratory Report 9, 1953.

Guilford, J.P., & Hoepfner, R. The analysis of intelligence. New York: McGraw-Hill, 1971.

Guilford, J.P., Kettner, N.W., & Christensen, P.R. A factor-analytic study across the domains of reasoning, creativity and evaluation. II. Administration of tests and analysis of results. University of Southern California Psychological Laboratory Report 16, 1956.

Guilford, J.P., & Lacey, J.I. (eds.). Printed classification tests. AAF Aviation Psychology Research Program Reports: Report 5. Washington, D.C.: G.P.O., 1947.

Hertzka, A.F., Guilford, J.P., Christensen, P.R., & Berger, R.M. A factor-analytic study of evaluative abilities. Educational and Psychological Measurement, 1954, 14, 581-597.

Hollingsworth, H.L. The definition of judgment. Psychological Review, 1925, 32, 337-361.

Johnson, D.M. A systematic treatment of judgment. Psychological Bulletin, 1945, 42, 193-224.

Johnson, D.M. The psychology of thought and judgment. New York: Harper, 1955.

Kettner, N.W., Guilford, J.P., & Christensen, P.R. A factor-analytic study across the domains of reasoning, creativity and evaluation. Psychological Monographs, 1959, 73, No. 479.

Nihira, K., Guilford, J.P., Hoepfner, R., & Merrifield, P.R. A factor analysis of the semantic evaluation abilities. University of Southern California Psychological Laboratory Report 32, 1964.

Pawlik, K. Concepts and calculations in human cognitive abilities. In Catell, R.B.(ed.), Handbook of multivariate experimental psychology. Chicago: Rand McNally, 1966, chap. 18.

Thurstone, L.L. Primary mental abilities. Psychometric Monographs, 1, 1938.

U.S. Civil Service Commission. Preparing for the Professional and Administrative Career Examination. Washington, D.C.: Author, 1974.

Vernon, P.E. The structure of human abilities. London: Methuen and Co., Ltd., 1950.

Wechsler, D. Measurement and appraisal of adult intelligence. Baltimore: The Williams and Wilkins Co., 1958.

ANNOTATED BIBLIOGRAPHY

Adkins, D.C., & Lyerly, S.B. Factor analysis of reasoning tests. Chapel Hill, N.C.: University of North Carolina Press, 1952.

This report presents the results of a re-analysis of 38 Air Force Tests as well as a detailed analysis of 66 tests of various types of reasoning. Examples of items from tests are given. Major factors extracted are discussed in detail. Extensive summary statistics (test means, test intercorrelations, factor loadings, etc.) are presented.

Balinski, B. An analysis of the mental factors of various age groups from nine to sixty. Genetic Psychology Monographs, 1941, 23, 191-234.

A factor analysis of the subscales of the Wechsler-Belleview I.Q. test is presented in this study. The cross-sectional analysis was repeated for six different age groups and changes in factor structure for the different groups is discussed in detail. The changes in importance and use of abilities with progressive age are given extensive attention.

Berger, R.M., Guilford, J.P., & Christensen, P.R. A factor-analytic study of planning abilities. Psychological Monographs, 1957, 71, No. 435.

This article reports on the same investigation presented in great detail by Guilford, Berger and Christensen in the University of Southern California Psychological Laboratory Reports 10 and 12.

Cattell, R.B. Abilities: their structure, growth and action. Boston: Houghton-Mifflin, 1971.

In this book Cattell presents his own views on the development and structure of human abilities. The theory of crystallized and fluid intelligence forms the basis for discussion of his model of ability structure. Judgment as a separate primary ability is given little attention since Cattell considers it a "narrower, less substantiated" Primary Mental Ability (PMA).

Corter, H.M. Factor analysis of some reasoning tests. Psychological Monographs, 1952, 66, No. 340.

Corter describes the reasoning process and reports on an analysis of a 21-test battery. A Judgment factor was extracted and it and the other factors obtained are discussed relative to their test content and meaning.

Dewey, J. How we think. New York: D.C. Heath, 1933.

An analysis of thought processes with emphasis on the training of thought via education is the subject of Dewey's landmark book. A chapter on the role of judgment in thinking activity discusses the factors important in and the functions of judgment.

French, J.W. The description of aptitude and achievement tests in terms of rotated factors. Psychometric Monographs, 5, 1951.

This monograph represents a summarization of all major factor-analytic studies carried out on batteries of aptitude and achievement tests, prior to 1951. Brief descriptions of all tests are presented, as well as complete tables of factors and loadings. Major factors are described and discussed and the tests with significant loadings on each are listed. Judgment is included.

Guilford, J.P. Structure of intellect factors and their tests. University of Southern California Psychological Laboratory Report 36, 1966.

This report presents an excellent concise description and explanation of the operations, contents, and products which comprise the structure of intellect model. Those factors which have been substantiated are listed and defined, along with the major tests which measure each one.

Guilford, J.P. The nature of human intelligence. New York: McGraw-Hill, 1967.

In this volume Guilford presents some historical background and general theories relative to the investigation of intelligence. He then explains, at length, his own structure of intellect model and discusses its implications for psychological theory. Finally, the physical and environmental determiners of intelligence and intellectual development and decline are reviewed.

Guilford, J.P., Berger, R.M., & Christensen, P.R. A factor-analytic study of planning. I. Hypotheses and description of tests. University of Southern California Psychological Laboratory Report, 10, 1954.

The various specific abilities required for successful planning activities are hypothesized and discussed in this report. Tests used in the investigation are described. The relationships of planning abilities to other already established mental abilities are also commented upon.

Guilford, J.P., Berger, R.M., & Christensen, P.R. A factor-analytic study of planning. II. Administration of tests and analysis of results. University of Southern California Psychological Laboratory Report 12, 1955.

This second part of the investigation of planning abilities describes the methodology used and presents a detailed analysis of the results. Fourteen factors were identified, of which Judgment was one. Factor loadings for all tests and factors are tabulated, and hypothesized factors are compared with those actually extracted.

Guilford, J.P., Comrey, A.L., Green, R.F., & Christensen, P.R. A factor-analytic study of reasoning abilities. I. Hypotheses and description of tests. University of Southern California Psychological Laboratory Report 1, 1950.

Reasoning hypotheses based on previous findings (Army Air Force factors) are elaborated on in this report. Individual descriptions of tests used in the investigations are presented.

Guilford, J.P., Green, R.F., & Christensen, P.R. A factor-analytic study of reasoning abilities. II. Administration of tests and analysis of results. University of Southern California Psychological Laboratory Report 3, 1951.

The methodology and detailed results of the study of reasoning abilities appear in this paper. Reasoning and its numerous components are defined, and differences between previous findings and those from this investigation are pointed out and discussed.

Guilford, J.P., Green, R.F., Christensen, P.R., Hertzka, A.F., & Kettner, N.W. A factor-analytic study of Navy reasoning tests with the Air Force Aircrew Classification Battery. University of Southern California Psychological Laboratory Report 6, 1952.

In this study the Judgment factor was well-defined and indicated a greater generality than in the Air Force investigations. Test descriptions, reliabilities and factor loadings are presented. Descriptions and discussion of factors are included, with reference to, and comparison with, their previous identifications.

Guilford, J.P., Green, R.F., Christensen, P.R., Hertzka, A.F., & Kettner, N.W. A factor-analytic study of Navy reasoning tests with the Air Force Aircrew Classification Battery. Educational and Psychological Measurement, 1954, 14, 301-325.

This journal article is a condensation of the methods, results and conclusions contained in the USC Psychological Laboratory Report 6, briefly reviewed above.

Guilford, J.P., Hertzka, A.F., Berger, R.M., & Christensen, P.R. A factor-analytic study of evaluative abilities. I. Hypotheses and description of tests. University of Southern California Psychological Laboratory Report 7, 1952.

Pertinent hypotheses relative to conceptual and perceptual evaluative abilities are presented and discussed in this report. The tests proposed to substantiate each hypothesis are described. The report includes a good review of the literature concerning the Judgment factor; one of the purposes of the study was to relate this factor to the broader concept of evaluation.

Guilford, J.P., Hertzka, A.F., & Christensen, P.R. A factor-analytic study of evaluative abilities. II. Administration of tests and analysis of results. University of Southern California Psychological Laboratory Report 9, 1953.

The methodology and statistical results of the study of evaluative abilities are compiled in this report. The obtained factors did not support some of the hypotheses presented in Report 7. The extracted factors are described and discussed. Although Judgment did not appear as a distinct factor, its possible relationship to other factors is analyzed.

Guilford, J.P., & Hoepfner, R. The analysis of intelligence. New York: McGraw-Hill, 1971.

This book presents a summary of all of the research done on the factors in the structure of intellect model and the abilities they represent. All of the tests used in the various studies carried out by the USC Psychological Laboratory are described and statistics relative to their factor loadings are presented. There is a good section on the evaluative abilities with discussion of their evolution, meaning, relation to other abilities and role in the overall intellective process.

Guilford, J.P., Kettner, N.W., & Christensen, P.R. A factor-analytic study across the domains of reasoning, creativity and evaluation. I. Hypotheses and description of tests. University of Southern California Psychological Laboratory Report 11, 1954.

Since many similar or related factors have appeared in separate studies of these three domains, this investigation concerns an examination of the nature of factors distinct and common to the three. A good literature review is presented.

Guilford, J.R., Kettner, N.W., & Christensen, P.R. A factor-analytic study across the domains of reasoning, creativity and evaluation. II. Administration of tests and analysis of results. University of Southern California Psychological Laboratory Report 16, 1956.

The result of this investigation was the extraction of twenty distinct factors. Nine, including Judgment, were essentially the same as those found in previous research. Factors are described and discussed but no new light is thrown on the separation of Judgment and Redefinition.

Guilford, J.P., & Lacey, J.I. (eds.). Printed classification tests. AAF Aviation Psychology Research Program Reports: Report 5. Washington, D.C.: G.P.O., 1947.

This book is a compilation of all research done by the Army Air Force Aviation Psychology Program (results are condensed and summarized in French, 1951). There are excellent chapters on Judgment and Planning which include examples of items in all tests used in these batteries, statistical analysis of the tests, and discussions of the relation of these factors to reasoning, perceptual and other abilities.

Guilford, J.P., Wilson, P.C., & Christensen, P.R. A factor-analytic study of creative thinking. II. Administration of tests and analysis of results. University of Southern California Psychological Laboratory Report 8, 1952.

This report summarizes the methodology and results of the study of Creative Thinking (hypotheses and description of tests appear in Report 4, 1951 by the same authors). A practical judgment measure was part of the test battery and it loaded on the Verbal, Visualization, General Reasoning and Redefinition factors. A conclusion was that the term "practical judgment" is not well understood.

Hertzka, A.F., Guilford, J.P., Christensen, P.R., & Berger, R.M. A factor-analytic study of evaluative abilities. Educational and Psychological Measurement, 1954, 14, 581-597.

In this study a Redefinition factor was found which was considered related to the Judgment factor. The Practical Judgment test loaded significantly (.42) on a General Reasoning factor.

Hoffman, K.I., Guilford, J.P., Hoepfner, R., & Doherty, W.J. A factor analysis of the figural cognition and figural evaluation abilities. University of Southern California Psychological Laboratory Report 40, 1968.

A good discussion of various definitions and components of evaluation is presented in this report.

Hogadone, E., & Smith, L.F. Some evidence on the validity of the Cardall Test of Practical Judgment. Journal of Applied Psychology, 1947, 31, 54-56.

This test consisted of multiple choice items similar to those in the Air Force Practical Judgment tests. The correlation between test scores and supervisory ratings of "practical judgment" were low (.24) and the Cardall Test was therefore considered not valid for measuring judgment ability. (The test has been deleted from Buros Mental Measurements Yearbook.)

Hollingsworth, H.L. The definition of judgment. Psychological Review, 1925, 32, 337-361.

This paper discusses the many aspects of judgment, both philosophical and practical. Psychophysical judgment is emphasized, but generalization to more abstract kinds of judgment is made. Hollingsworth calls judgment a "process" and explains what it is, as well as what is isn't.

Johnson, D.M. A systematic treatment of judgment. Psychological Bulletin, 1945, 42, 193-224.

This paper presents another excellent discussion of judgment - its general principles, complexity and the different types of judgments which can be made depending upon situational requirements.

Johnson, D.M. The Psychology of thought and judgment. New York: Harper, 1955.

Johnson has written a very complete and readable account of judgment and its relation to thought. The book contains discussions of what good judgment is and what it means to define and measure judgment as an ability. Research on the Judgment factor and the tests which have been used are reviewed.

Kettner, N.W., Guilford, J.P., & Christensen, P.R. A factor-analytic study across the domains of reasoning, creativity and evaluation. Psychological Monographs, 1959, 73, No. 479.

This journal article presents the same material appearing in USC Psychological Laboratory Reports 11 and 16, (this monograph may be more readily available in libraries).

Mandell, M. The Administrative Judgment Test. Journal of Applied Psychology, 1950, 34, 145-147.

Mandell reports on a test developed by the U.S. Civil Service Commission with items similar to the type appearing in the Air Force Tests. This test is totally job-oriented, however. The AJT was correlated with vocabulary (.59) and with collective supervisor ratings (.60).

McKillip, R., Trattner, M., Corts, D., & Wing, H. Professional and Administrative Career Examination: Research and Development. Manuscript in preparation, Washington, D.C.: U.S. Civil Service Commission, 1976.

The research (including job analysis and literature review) conducted for the purpose of developing the PACE is presented in this report. Ability constructs necessary for the performance of the 27 jobs included in the analysis and the measurement of these abilities are explained and discussed.

Munoz-Colberg, M. Background of experimental questions for testing judgment. Operations Paper. Washington, D.C.: U.S. Civil Service Commission, 1975.

This paper succinctly summarizes the intellective and content requirements for items measuring judgment. The distinction between Reading Comprehension, Inference and Comprehension (Judgment) item types is emphasized.

Nihira, K., Guilford, J.P., Hoepfner, R., & Merrifield, P.R. A factor analysis of the semantic evaluation abilities. University of Southern California Psychological Laboratory Report 32, 1964.

This is a very comprehensive report on a study of the semantic evaluation abilities. It contains background information on judgment - its role in problem solving, and the separation of cognition and evaluation. Semantic evaluation factors, as well as all tests to measure them, are described. Reference tests for related factors (e.g., Cognition) are included. There are complete tables of test intercorrelations and factor loadings. This report was extremely helpful during the search for new item types for measuring judgment.

Pawlik, K. Concepts and calculations in human cognitive abilities. In Cattell, R.B. (ed.), Handbook of multivariate experimental psychology. Chicago: Rand McNally, 1966, chap. 18.

Pawlik has written a very good analysis of the nature and structure of abilities and of the investigations carried out to identify and define them. He discusses, at length, broad intellectual factors (e.g., Perception and Visual Thinking, Verbal Fluency and Flexibility of Thinking, Problem Solving) and places specific abilities (e.g., deduction, reasoning, judgmental processes) in these broad categories. Pertinent research studies are cited. Pawlik concludes with a good description of Cattell's theory of intelligence (crystallized and fluid).

Thurstone, L.L. Primary mental abilities. Psychometric Monographs, 1, 1938.

This is Thurstone's classic monograph in which he presents his factor-analytic research on the identification of primary mental abilities. The significance of factors and factor loadings are elaborated. A distinct Judgment factor was not identified by Thurstone.

Vernon, P.E. The structure of human abilities. London: Methuen and C., Ltd., 1950.

This book includes chapters on factor analysis (its limits and uses) and on the numerous theories of the structure of abilities. Differential aptitude tests are reviewed, including the U.S. Employment Service's General Aptitude Test Battery and Army Air Force Classification Batteries from World War II. The British hierarchical structure of abilities is emphasized and is compared and contrasted with American findings.

Wechsler, D. Measurement and appraisal of adult intelligence. Baltimore: The Williams and Wilkins Co., 1958.

The Wechsler Adult Intelligence Scale is the subject of careful and detailed description and analysis in this volume. The selection and description of the subscales are discussed. Factor-analytic studies of the subscales are presented as well as research on the changes in intelligence with reference to age and sex. Diagnostic and practical applications are also given attention.